

Requested document:	<a href="#">JP2000165831 click here to view the pdf document</a>
---------------------	--

## MULTI-POINT VIDEO CONFERENCE SYSTEM

Patent Number:

Publication date: 2000-06-16

Inventor(s): SATODA KOZO; HIRAIKE RYUICHI

Applicant(s): NIPPON ELECTRIC CO

Requested Patent: ☐ [JP2000165831](#)

Application Number: JP19980339131 19981130

Priority Number(s): JP19980339131 19981130

IPC Classification: H04N7/15

EC Classification: [H04N7/14A2](#), [H04N7/14A3](#), [H04N7/15](#)

Equivalents: ☐ [US6285392](#)

---

### Abstract

PROBLEM TO BE SOLVED: To provide natural video and audio data with high presence to each conference terminal participant in a multi-point video conference system. SOLUTION: A plurality of cameras are connected to conference terminals 1i (i=1,...,n) placed at remote points to photograph video images of participants. A multi-point video input section 21 of a central controller 2 receives video images obtained by photographing the participants in multi-directions from multi-points. A video selection section 22 selects the video images of the participants by taking a direction into account in which how each participant sees the other conference participants in a virtual conference space, synthesizes them in the virtual space and each conference terminal exhibits a computer graphic video image on each conference terminal.

---

Data supplied from the **esp@cenet** database - I2

(51)Int.Cl.<sup>7</sup>識別記号F Iテーマコード\* (参考)  
H 0 4 N 7/156 1 0H 0 4 N 7/156 1 05 C 0 6 4

審査請求 有 請求項の数 7 O L （全 9 頁）

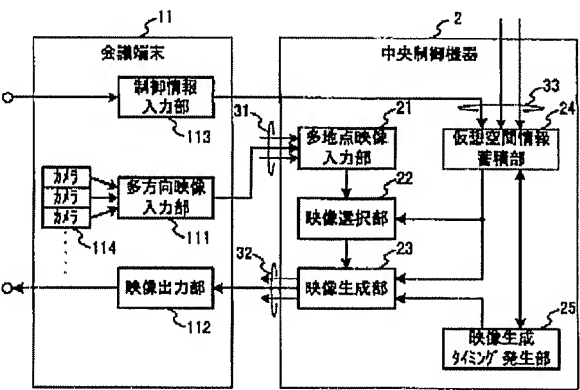
(21)出願番号	特願平10-339131	(71)出願人	000004237 日本電気株式会社 東京都港区芝五丁目7番1号
(22)出願日	平成10年11月30日 (1998. 11. 30)	(72)発明者	里田 浩三 東京都港区芝五丁目7番1号 日本電気株式会社内
		(72)発明者	平池 龍一 東京都港区芝五丁目7番1号 日本電気株式会社内
		(74)代理人	100105511 弁理士 鈴木 康夫（外1名） Fターム(参考) 5C064 AA02 AC02 AC06 AC08 AC14 AC16

(54)【発明の名称】 多地点テレビ会議システム

(57)【要約】

【課題】 多地点テレビ会議システムにおいて、各会議端末参加者に臨場感の高い 自然な映像と音像を提供する。

【解決手段】 遠隔地にある n 地点の各会議端末 1 i（i = 1, . . . , n）に複数のカメラを接続し参加者の映像撮影する。中央制御機器 2 では多地点からの多方向で参加者を撮影した映像を多地点映像入力部 21 で受信する。各会議端末の参加者は仮想的な会議空間内で他の会議参加者がどのように見えるかという方向を考慮し、映像選択部 22 で参加者の映像を選択し、それを仮想空間に合成し、コンピュータグラフィックス映像を各会議端末で、参加者に提示する。



【特許請求の範囲】

【請求項1】 会議を制御する中央制御機器と、前記中央制御機器と接続された複数の会議端末から成る多地点テレビ会議システムにおいて、

前記会議端末は、該会議端末を介して会議に参加する参加者を複数の方向から撮影する複数の撮像装置と、該複数の撮像装置の映像を入力して前記中央制御機器に前記映像を送信する多方向映像入力部と、前記中央制御機器から送信されてきた仮想的な会議室の映像を受信して前記参加者に提示する映像出力部と、前記参加者から会議を制御するための情報を受信して中央制御機器へ送信する制御情報入力部とを備え、

前記中央制御機器は、前記各参加者の会議端末から送信されてきた映像を受信する多地点映像入力部と、前記各参加者の会議端末に映像を合成して送信するために、どの参加者が見る映像を中央制御機器で合成して送信するかを指示する映像生成タイミング発生部と、前記仮想的な会議空間の映像を生成するための情報を蓄えておき、前記各会議端末からの制御情報により会議空間の情報を更新するとともに、前記映像生成タイミング発生部からの指示により映像を生成するための情報を送信する仮想空間情報蓄積部と、前記多地点映像入力部で受信した前記多方向の映像のうち前記仮想空間情報蓄積部からの情報により、現在映像を生成している対象の参加者が前記仮想的な会議空間内で他の会議端末の参加者を見た場合の見え方を考慮することによって前記多地点映像入力部で受信した前記多方向の映像の中から適切な映像を選択する映像選択部と、前記映像選択部で選択された映像を前記仮想空間情報蓄積部からの情報により前記仮想的な会議空間に配置した映像を生成して各参加者の会議端末に送信する映像生成部とを備えていることを特徴とする多地点テレビ会議システム。

【請求項2】 前記中央制御機器に備えられている映像選択部は、前記多地点映像入力部で受信した前記多方向の映像の中に適切な映像がない場合に、前記他方向の映像の中の複数の映像を用いて適切な映像を生成する機能を有していることを特徴とする請求項1記載の多地点テレビ会議システム。

【請求項3】 前記会議端末は、会議参加者の音声を入力して前記中央制御機器に送信する音声入力部と、前記中央制御機器で合成された音声を受信し参加者に提示する音声出力部と、参加者から会議を制御するための情報を受信して前記中央制御機器へ送信する制御情報入力部とを備え、

前記中央制御機器は、前記仮想的な会議空間の音声を生成するための情報を蓄えておき、各会議端末からの制御情報により会議空間の情報を更新し、音声を生成するための情報を送信する仮想空間情報蓄積部と、各参加者の会議端末から送信されてきた音声を受信する多地点音声入力部と、多地点多方向の音声のうち前記仮想空間情報

蓄積部からの情報により、音声を生成している対象の参加者が仮想的な会議空間内で他の会議端末の参加者の音声の聞こえ方を考慮することによって適切に音声を生成し前記会議端末に送信する映像合成部とを備えていることを特徴とする請求項1記載の多地点テレビ会議システム。

【請求項4】 会議に参加する参加者を複数の方向から撮影する複数の撮像装置と、該撮影された複数の撮像装置の映像を入力して中央制御機器に前記映像を送信する多方向映像入力部と、前記中央制御機器から送信されてきた仮想的な会議室の映像を受信して前記参加者に提示する映像出力部と、前記参加者から会議を制御するための情報を受信して前記中央制御機器へ送信する制御情報入力部とを備えていることを特徴とする多地点テレビ会議システム用端末。

【請求項5】 複数の会議端末からそれぞれ送信されてきた映像を受信する多地点映像入力部と、前記複数の会議端末に映像を合成して送信するために、どの会議端末の参加者が見る映像を合成して送信するかを指示する映像生成タイミング発生部と、仮想的な会議空間の映像を生成するための情報を蓄えておき、前記各会議端末から受信した制御情報により会議空間の情報を更新するとともに、前記映像生成タイミング発生部からの指示により映像を生成するための情報を送信する仮想空間情報蓄積部と、前記多地点映像入力部で受信した前記多方向の映像のうち前記仮想空間情報蓄積部からの情報により、現在映像を生成している対象の参加者が前記仮想的な会議空間内で他の会議端末の参加者を見た場合の見え方を考慮することによって前記多地点映像入力部で受信した前記多方向の映像の中から適切な映像を選択する映像選択部と、前記映像選択部で選択された映像を前記仮想空間情報蓄積部からの情報により前記仮想的な会議空間に配置した映像を生成して各参加者の会議端末に送信する映像生成部とを備えていることを特徴とする多地点テレビ会議システム用中央制御機器。

【請求項6】 異なる地点に配置された複数の会議端末を互いに接続して成る多地点テレビ会議システムにおいて、

前記各会議端末は、それぞれ、各会議端末を介して会議に参加する参加者を複数の方向から撮影する複数の撮像装置と、該撮影された複数の撮像装置の映像を入力して前記接続されている複数の会議端末に送信する多方向映像入力部と、前記参加者から会議を制御するための情報を受信して前記複数の会議端末へ送信する制御情報入力部と、前記複数の会議端末から送信されてきた映像を受信する多地点映像入力部と、仮想的な会議空間の映像を生成するための情報を蓄えておき、前記複数の会議端末から受信した前記制御情報により会議空間の情報を更新し、映像を生成するための情報を出力する仮想空間情報蓄積部と、前記仮想空間情報蓄積部からの情報により、

自端末の参加者が仮想的な会議空間内で他の会議端末の参加者を見た場合の見え方を考慮することによって前記複数の会議端末から受信した多地点多方向の映像のうち適切な映像を選択する映像選択部と、前記映像選択部で選択された映像を前記仮想空間情報蓄積部からの情報により仮想的な会議空間に配置して前記複数の会議端末の各参加者が仮想的な会議空間を共有している映像を生成する映像生成部とを備えていることを特徴とする多地点テレビ会議システム。

【請求項7】 会議に参加する参加者を複数の方向から撮影する複数の撮像装置と、該撮影された複数の撮像装置の映像を入力して他の会議端末に送信する多方向映像入力部と、前記参加者から会議を制御するための情報を受信して前記他の会議端末へ送信する制御情報入力部と、前記他の会議端末から送信されてきた映像を受信する多地点映像入力部と、仮想的な会議空間の映像を生成するための情報を蓄えておき、前記他の会議端末から受信した前記制御情報により会議空間の情報を更新し、映像を生成するための情報を出力する仮想空間情報蓄積部と、前記仮想空間情報蓄積部からの情報により、自端末の参加者が仮想的な会議空間内で前記他の会議端末の参加者を見た場合の見え方を考慮することによって前記他の会議端末から受信した多地点多方向の映像のうち適切な映像を選択する映像選択部と、前記映像選択部で選択された映像を前記仮想空間情報蓄積部からの情報により仮想的な会議空間に配置して前記他の会議端末の参加者が仮想的な会議空間を共有している映像を生成する映像生成部とを備えていることを特徴とする多地点テレビ会議システム用端末。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、多地点テレビ会議用映像および音声合成装置に係り、特に臨場感の高い多地点テレビ会議用映像および音声合成装置に関する。

【0002】

【従来の技術】従来、多地点テレビ会議システムにおいて、参加者に他の地点の参加者の映像を提示する手法として、図10に示すように、1台のモニタに他の地点の映像1B、1C、1Dをマルチウィンドウで表示したり、複数のモニタにそれぞれ1Bの映像、1Cの映像、1Dの映像を表示するという手法が実現されている。しかしながら、この手法は現在の話者が誰であるのかわかりにくく、話者の視線と話を聞いている人の視線が一致せず会話の状態が不自然であるという欠点があった。この解決手法として、特開平6-351013号公報で提案されている複数カメラを用いる手法がある。この手法は図11で示すように、モニタの中で各参加者の映像1B、1C、1Dが映っている場所にそれぞれカメラ114B、114C、114Dを設置する。そして、114Bのカメラの映像は1Bの参加者の端末に送信す

る。同様に114C、114Dのカメラも1C、1Dへ送信し、他の参加者のカメラも同様の配置をする。

【0003】このように配置すると、例えば1Aが1Bの映像に向かって会話をしているときは、1Bには1Aの正面映像が映り、1C、1Dには1Aの横からの映像が映ることになる。この手法によれば、視線の一致による会議の臨場感を増すことができるが、自分を含めて $n$ 地点の参加者がいる場合には、各参加者の端末にそれぞれ $(n-1)$ 台のカメラを設置する必要があり、参加者が増えると設置が困難となる。

【0004】また、さらに臨場感を高めるため、特開平7-67035号公報で提案されているように、コンピュータグラフィックス(CG)を用いて仮想的な会議空間の映像を生成し参加者に提示する手法もある。これは、図12で示すように、机やその他会議室の背景に各参加者の映像を配置することにより、さらに会議に参加しているという一体感を持たせることができる。この手法では、特開平6-351013号公報で提案されている複数カメラによる視線一致の手法も用いており、同様に参加者が増えると設置が困難となる。

【0005】

【発明が解決しようとする課題】このように、従来のテレビ会議システムでは、会議をしている臨場感を出すために、主に話者と視聴者の間での視線を一致させることにより、会議に参加している臨場感を出すことを目的としているため、 $n$ 人参加の会議では、各端末に $(n-1)$ 台のカメラを設置する必要があり、実現は容易でなかった。

【0006】本発明の目的は、各端末に参加者の人数以下の数台のカメラを設置し、コンピュータグラフィックス上に参加者の映像を合成することによって、少ない台数のカメラを用いて臨場感の高い仮想的な会議空間の映像を作成する手段を提供することにある。

【0007】

【課題を解決するための手段】本発明は、上記の目的を達成するため、第1の様態として 以下のような機器を有する多地点会議用映像音声合成装置を特徴とする。

【0008】まず、全体的構成としては、中央で会議を制御する中央制御機器と、この中央制御機器と接続され、各参加者が会議に参加するための会議端末から成っている。この会議端末は、各地点参加者ごとに用意されており、参加者を複数の方向から撮影する撮像装置と、この複数の撮像装置の映像を入力し、中央制御機器に映像を出力する多方向映像入力部と、中央制御機器から送信されてきた仮想的な会議室の映像を受信し、参加者に提示する映像出力部と、参加者から映像以外の会議を制御するための情報、例えば会議参加や途中退室、資料提示などの情報を受信し、中央制御機器へ送信する制御情報入力部から成る。

【0009】また、中央制御機器は、各参加者の会議端

末から送信されてきた映像を受信する多地点映像入力部と、各参加者の会議端末に映像を合成して送信するために、どの参加者が見る映像を中央制御機器で合成して送信するかを指示する映像生成タイミング発生部と、仮想的な会議空間の映像を生成するためのコンピュータグラフィックスの3次元モデルや、各参加者が仮想空間のどこに配置されるかなどの情報を蓄えておき、各会議端末からの制御情報により会議空間の情報を更新し、また映像生成タイミング発生部からの指示により映像を生成するための情報を送信する仮想空間情報蓄積部と、多地点多方向の映像のうち仮想空間情報蓄積部からの情報により、現在映像を生成している対象の参加者が仮想的な会議空間内で他のユーザを見るとどのように見えるかを考慮することによって適切な映像を選択する映像選択部と、映像選択部で選択された映像を仮想空間情報蓄積部からの3次元モデルなどの情報により仮想的な会議空間に配置し、各参加者が仮想的な会議空間を共有している映像を生成して各参加者の会議端末に送信する映像生成部とから成る。

【0010】また、本発明は、第二の様態として、第一の様態の映像選択部のかわりとして、映像を多地点多方向映像の映像を選択する時に適切な映像がない場合、複数の映像から適切な映像を生成することができる映像選択生成部を備えた中央制御装置からなる。また、本発明の第三の様態は、映像だけでなく音声によっても臨場感を得るために、会議参加者の音声を入力して中央制御機器に送信する音声入力部と、中央制御機器で合成された音声を受信し、参加者に提示する音声出力部と、会議を制御するための情報を受信し、中央制御機器へ送信する制御情報入力部とから成る会議端末と、各参加者の会議端末から送信されてきた音声を受信する多地点音声入力部と、各参加者が仮想空間のどこに配置されるかなどの情報を蓄えておき、各会議端末からの制御情報により会議空間の情報を更新し、映像を生成するための情報を送信する仮想空間情報蓄積部と、多地点多方向の音声のうち仮想空間情報蓄積部からの参加者の配置情報により、音声を生成している対象の参加者が仮想的な会議空間内で他のユーザからの音声がどのように聞こえるかを考慮することによって適切に音声を生成し、前記会議端末に送信する音声合成部とから成る中央制御部とを備えた多地点会議用映像音声合成装置を特徴とする。

【0011】本発明の第一の様態では、中央制御機器と各参加者端末をそれぞれ結んでいるが、各参加者端末間をすべて結び、各参加者端末において各自の映像を生成することにより、中央制御機器ですべての参加者端末の映像を生成するための時間差を減らすことができる。

【0012】このため、本発明の第四の様態は、各参加者を複数の方向から撮影する撮像装置と、複数の撮像装置の映像を入力し、各参加者の会議端末に映像を出力する多方向映像入力部と、会議を制御するための情報を受

信し各参加者の会議端末に送信する制御情報入力部と各参加者の会議端末から送信されてきた映像を受信する多地点映像入力部と、仮想的な会議空間の映像を生成するための情報を蓄えておき、各会議端末からの制御情報により会議空間の情報を更新し、映像を生成するための情報を送信する仮想空間情報蓄積部と、多地点多方向の映像のうち前記仮想空間情報蓄積部からの情報により、現在自分の端末の参加者が仮想的な会議空間内で他のユーザを見るとどのように見えるかを考慮することによって適切な映像を選択する映像選択部と、映像選択部で選択された映像を仮想空間情報蓄積部からの情報により仮想的な会議空間に配置し各参加者が仮想的な会議空間を共有している映像を生成する映像生成部とからなる会議端末がそれぞれ接続されている多地点会議用映像音声合成装置を特徴とする。

【0013】

【発明の実施の形態】以下、図面を参照しながら、本発明の実施の形態を詳細に説明する。図1は、本発明の第1の実施の形態に係る多地点会議用映像音声合成装置の構成を示すブロック図である。

【0014】この多地点会議用映像音声合成装置は、中央で会議の制御を行う中央制御機器2と、遠隔地から会議に参加する会議端末からなっている。ここでは4地点の参加者が多地点会議に参加し、そのそれぞれが会議端末1*i* (*i* = 1, 2, 3, 4) を中央制御機器2にネットワークを介して接続している状態を考える。図1では、会議端末11のみを示しているが、会議端末12, 13, 14も同様の構成を有している。なお、これを*n*地点に拡張するのは容易である。

【0015】会議端末11は、複数のカメラ114を接続し、参加者の映像を多方向から撮影する。カメラは少なくとも3台用意し、参加者の正面、左前、右前から撮影されるように配置する。カメラの配置は必ずしも正確に配置されていなくてもよいが、正確に測定して配置しておけば、より臨場感の高い映像を生成することができる。

【0016】参加者を複数の方向から撮影したカメラ群114の映像は、多方向映像入力部111に入力され、さらにこの映像は中央制御機器2に送られるが、その際、複数方向の映像を時分割で送信しても良いし、あるいは、図5で示すように、一枚の画面情報を、4分割、9分割、16分割などに分割して一画面の中に複数の映像を合成してもよい。このようにして合成された映像は中央制御機器2へ送られる。

【0017】中央制御機器2では、各会議端末1*i* (*i* = 1, 2, 3, 4) の多方向映像入力部111からの入力を多地点映像入力部21で受信する。また、中央制御機器2では、仮想的な会議室の映像を生成して各会議端末1*i*へ送信する。この仮想的な会議室は、例えば、机や会議室内部等はCGの3次元形状モデルからできてお

り、その中に、机を囲んで各端末の参加者の映像を配置したものである。このような会議室の机や会議室内部の3次元形状モデルや各参加者を仮想的な会議室のどこに配置するかという仮想空間に関する情報は、仮想空間情報蓄積部24により管理する。ここでの説明では、図6に示されているように、4人の参加者が会議テーブルを囲んで会議を行っているという仮想空間を想定する。

【0018】各参加者は仮想的な会議空間内に存在して会議に参加しているので、各参加者が仮想空間内のそれぞれの位置に配置される場合、各参加者は自分の配置されている場所から仮想的な会議空間を眺めているのが自然である。従って、図6の配置の場合、1Aの参加者から見ると、正面に1Cが見え、右前に1Dが、左前に1Bが見える。一方、1Bから見ると、正面は1Dで、右前に1A、左前に1Cが見えているはずである。

【0019】すなわち、各参加者端末において参加者が見る仮想的な会議空間の映像は、各会議端末によって異なる。中央制御機器2では、この異なる映像すべてを生成して各会議端末に送信する必要がある。このために、中央制御機器2では、現在どの会議端末用の映像を生成しているかのタイミングを生成する映像生成タイミング発生部25を用意する。映像生成タイミング発生部25は、現在の映像を生成しなければならない会議端末のIDを仮想空間情報蓄積部24に送信する。

【0020】仮想空間情報蓄積部24は、その会議端末のIDから、現在生成している仮想空間の映像の視点位置、視線方向、その位置から見た他の会議参加者の相対位置の情報を映像選択部22と映像生成部23に送信する。例えば、映像生成タイミング発生部25が、現在1Aの映像を生成しているというタイミング情報を流せば、仮想空間情報蓄積部24は、仮想空間映像を生成するための視点の位置として1Aの視点場所を指定し、1Bが左前、1Cが正面、1Dが右前に見えるという情報を映像選択部22と映像生成部23に送信する。

【0021】映像選択部22では、多地点映像入力部21で受信した多地点での参加者を多方向から撮影したそれぞれの映像から、1Bについては右前から撮影したカメラ映像を、1Cについては正面から撮影したカメラ映像を、1Dについては左前から撮影したカメラ映像を選択し、映像生成部23に送信する。映像生成部23では、仮想空間情報蓄積部24から送信されてきた会議室の3次元形状モデルに、映像選択部22で選択した各参加者の映像を仮想空間に配置し、1Aの視点で仮想的な会議空間の映像を生成する。そして、映像生成タイミング発生部25で設定されているタイミングで1Aの会議端末へ映像を送信する。

【0022】映像生成タイミング発生部25は、1Aへの映像送信後、次は1Bの映像を生成するように次のタイミングへ移行する。このように中央制御機器2では、この映像生成をすべての参加者に対して行い、それぞれ

の端末へ送信する。各会議端末1iは、この自分に送信されてきた映像を映像出力部1i2で受け取り、その端末の利用者に映像を提示する。また、制御情報入力部1i3は、映像情報以外の情報を中央制御機器2に送付する。

【0023】たとえば、その会議端末から新たに会議に参加する場合、その端末からの参加要請を中央制御機器2の仮想空間情報蓄積部24が受け取り、仮想空間情報蓄積部24は、新たな参加者を仮想空間内の適切な位置に配置する。そして、映像生成タイミング発生部25に新規参加者の存在を伝えることで、新規参加者の映像が仮想的な会議空間に配置され、その参加者も会議空間の映像を見ることができるようになる。

【0024】ここで、さらに会議空間に参加しているという臨場感を出すために、映像選択部22で選択された映像の背景を除去し、背景部分は透明にして仮想的な会議空間に参加者の映像だけをなじませれば、臨場感はより高くなる。この背景を除去する方法としては、背景を青い特殊な布などで覆い色情報をもとに背景を除去したり、背景の映像を予め撮影しておき、それとの差分をとることで参加者の映像だけを抽出したり、あるいは、複数のカメラをもちいたステレオ法などにより参加者の部分だけを抽出する等、様々な方法がある。

【0025】図7は、より臨場感を出すための映像配置の例を示しており、図12にも示されているように、従来の手法では、3次元形状モデルの机に参加者の映像を配置する場合、正面の映像はそのまま配置しても違和感を感じないが、左右に配置されて、机に座っている映像を出すために図7(a)のように正面画像を傾けて配置しているが、このようにすると、参加者の映像の画角が狭くなり、参加者映像が細く映ってしまい現実的でない映像になってしまう。

【0026】本発明では、図7(b)のように横方向からの映像をそのまま視線方向に垂直に配置することで画面を見ても自然な映像を生成することができる。また、これ以外にもカメラの設置台数が少ない場合等、正確な場所からのカメラ映像が得られない場合には、参加者の映像を傾けて配置する従来の技法を組み合わせることで、擬似的に視線一致の映像を生成することも可能である。

【0027】また、会議形態として、図6のような円卓を囲んだ会議から、プレゼンテーションがメインの会議など様々な会議形態があるが、仮想空間情報蓄積部24に種々の会議室の形状モデルを蓄えておき、会議形態に応じて3次元モデルを変更することにより会議の雰囲気を変えることができる。また、それに伴う参加者の配置の調整も容易に行うことができる。

【0028】本発明の第一の実施の形態においては、参加者1Aが見る映像を生成するために1B、1C、1Dの映像から、正面、あるいは横方向からの映像を選択し



ている。しかしながら、このような構成により、完全に視線を一致して臨場感のある映像を生成するためには、1 Aの端末では、1 Bの映像を配置する位置に1 B用のカメラを置く必要があり、参加者が増えた場合や会議の雰囲気によってレイアウトを変更するなどの状況には対応できない。第一の実施の形態では、その場合には、従来の映像を傾けて合成する技法も組み合わせることにより、擬似的に視線一致を行う手法を提案しているが、正確に実現することは困難である。

【0029】図2は、本発明の第二の実施の形態を示すブロック図である。本実施の形態では、第一の実施の形態における映像選択部22のかわりに映像選択生成部24を用意する。この映像選択生成部26も映像選択部22と同様に適切な映像を選択するが、さらに映像選択生成部26は、ちょうどカメラの配置にあうものがない場合に、会議端末から入力された異なる方向からの映像に基づいて所望の方向の映像を生成して出力する機能を有している。

【0030】例えば、図8のように、カメラ1144から1 Aを撮影した映像が欲しい場合に、カメラ1141および1143の両カメラの映像11411および11431しかない場合がありうる。その場合、この両方の映像から映像11441を生成するのが映像選択生成部26の動作である。具体的には、ちょうど良い場所からの映像がない場合、近くにある2つの映像を選択し、その間を映像のモーフィングなどの手法によって必要な映像を生成する。

【0031】その場合、会議端末1 iでのカメラの配置が正確にわかっていれば、仮想空間情報蓄積部24にそのカメラ配置情報を持つことにより、映像を変形するための正確なパラメータを計算し、臨場感の高い映像を得ることができる。しかし、そのような正確な配置は一般には困難なので、ある程度、正面、左前45度、右前45度といったパラメータを設定しておくことで、映像を生成することも可能である。

【0032】図3は、本発明の第三の実施の形態を示すブロック図である。今まで述べた実施例は、すべて臨場感の高い映像を生成するための説明であるが、実際1 Aから見た映像を生成し、左から右へ1 B、1 C、1 Dと配置されている場合は、1 Bの発言が左方から、1 Cは正面から、1 Dは右方から聞こえるのが自然である。そこで、本実施の形態では、音声についても該当する映像の方から聞こえるようにしている。

【0033】第三の実施の形態では、各会議端末1 iにマイク1 i7があり、マイク1 i7で集音された音声信号は、音声入力部1 i8に入力され、中央制御機器2に送信される。中央制御機器2では、多地点音声入力部26が各端末の音声情報を受け取る。音声合成部27では、多地点音声入力部26で入力した各地点の音声と仮想空間情報蓄積部24で蓄えられている参加者の配置情

報をもとにして、音声合成部27で会議端末1 iで聞かれる映像を生成し会議端末1 iへ送る。会議端末1 iでは、音声出力部1 i9が音声を受け取り参加者へ音声を提示する。

【0034】具体的には、図6のような配置の場合、参加者1 Aが聞く映像を合成するためには、1 Bの音は右の音を左より小さくし、1 Dはその逆に左の音を右より小さくする。1 Cは1 B、1 Dより離れているので、両方の音のレベルを少し下げて3地点の音をステレオに合成する。これを1 Aに送るとその場所に会議参加者がいるような臨場感を与えることができる。

【0035】以上で述べた実施の形態においては、映像、映像の生成を中央制御機器2で行っている。この構成の場合、会議への参加者が増えると中央制御機器2の負担が大きくなり、一つの中央制御機器2ですべての会議参加者の映像、映像の生成を行うのが困難になってくる。これに対処する一方法としては、中央制御機器2を複数にして、それぞれの会議参加者の映像、映像の生成を複数の中央制御機器で分担することにより、処理の低下を抑えることができる。

【0036】また、別の方法として、すべてのクライアントをそれぞれ接続する手法が考えられる。図4は、本発明の第四の実施の形態を示すブロック図である。本実施の形態においては、各会議端末に中央制御機器2の機能を分散したことを特徴としている。

【0037】図4において、会議端末11は、第一の実施の形態のように、複数のカメラ114の多方向映像入力部111により当該端末における参加者の映像を撮影し他のすべてのクライアントに送信する。一方、他のすべてのクライアントからの映像は、多地点映像入力部1110に入力され、仮想空間情報蓄積部1113に蓄えられている各参加者の配置情報により映像選択部1111で必要な映像を選択し、映像生成部1112で仮想的な会議空間の映像を生成する。この映像生成手順は第一の実施の形態で示した通りである。

【0038】また、本実施の形態の会議端末11に対して、図3に対応する音声合成手段を付加することにより、音声についても同様にして仮想的な会議空間の映像を各会議端末において生成することは容易である。

【0039】本実施の形態を実現するためのネットワーク結線は、図9(b)に示すようになる。図9は5地点での結線を示したものである。図9(a)で示すように、中央制御機器2に接続する場合は結線数が5である。一方、第四の実施の形態の場合は、図9(b)のように結線数は10になる。このように、全参加者端末を結ぶ方式はネットワーク結線が複雑になるが、ネットワーク回線の容量が十分にある場合には、機器構成の簡便さから有利である。一方、ネットワーク容量に比して中央制御機器2の映像生成能力が高い場合は、中央制御機器2に接続する中央結線方式の方が優れている。

【0040】

【発明の効果】本発明の第一、第二の実施の形態によれば、他の会議参加者の映像をコンピュータグラフィックスの仮想的な会議空間に配置することで視線の一致した自然な会議空間の映像を会議参加者に提示することができ、臨場感豊かな映像生成装置を提供できる。

【0041】本発明の第三の実施の形態によれば、映像だけでなく、音声も会議参加者の位置に応じて定位して合成することにより、臨場感豊かな音像生成装置を提供できる。

【0042】本発明の第四の実施の形態によれば、各会議端末において、映像、音像を合成することにより、小規模であるが簡単な多地点テレビ会議システムを構築することができる。

【図面の簡単な説明】

【図1】本発明の第一の実施の形態の構成を示すブロック図である。

【図2】本発明の第二の実施の形態の構成を示すブロック図である。

【図3】本発明の第三の実施の形態の構成を示すブロック図である。

【図4】本発明の第四の実施の形態の構成を示すブロック図である。

【図5】複数方向からの映像を一画面に多重化して送信する多方向映像入力部の説明図である。

【図6】4人参加の会議における仮想的な会議空間での参加者の配置の例である。

【図7】横方向にいる参加者の映像を自然に配置するための説明図である。

【図8】適切なカメラ位置がない場合、必要な方向からの映像を生成する方法の説明図である。

【図9】会議端末間の接続形態を説明するための図である。

【図10】従来のテレビ会議における映像提示法の例である。

【図11】従来のテレビ会議における視線一致を実現した映像提示法の例である。

【図12】従来のテレビ会議におけるコンピュータグラフィックスで映像を提示する方法の例である。

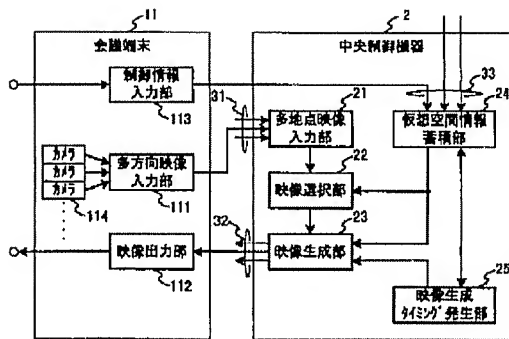
【符号の説明】

11 会議端末1  
111 会議端末1内の多方向映像入力部  
112 会議端末1内の映像出力部  
113 会議端末1内の制御情報入力部  
114 会議端末1内の多方向から参加者を撮影するカメラ群  
117 会議端末1内のマイク

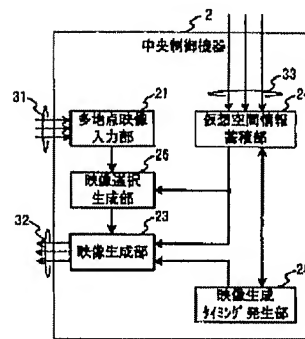
118 会議端末1内の音声入力部  
119 会議端末1内の音声出力部  
1110 会議端末1内の多地点映像入力部  
1111 会議端末1内の映像選択部  
1112 会議端末1内の映像生成部  
1113 会議端末1内の仮想空間情報蓄積部  
1141 会議端末1で撮影する正面カメラ  
1142 会議端末1で撮影する右前カメラ  
1143 会議端末1で撮影する左前カメラ  
11411 会議端末1で撮影する正面映像  
11421 会議端末1で撮影する右前映像  
11431 会議端末1で撮影する左前映像  
1111 会議端末1で複数方向映像を1画面内に多重化した映像の例  
1A 会議参加者1  
1B 会議参加者2  
1C 会議参加者3  
1D 会議参加者4  
1B1 会議参加者2の正面映像を机にあわせて配置した例  
1B2 会議参加者2の横向き映像を視線あわせて配置した例  
1144 会議端末1で欲しい方向の映像であるがカメラを設置していない場所  
11441 会議端末1で欲しい方向の映像を合成した映像  
114B 会議端末1で1Bへ送信するための映像を撮影するカメラ  
114C 会議端末1で1Cへ送信するための映像を撮影するカメラ  
114D 会議端末1で1Dへ送信するための映像を撮影するカメラ  
2 中央制御機器  
21 多地点映像入力部  
22 映像選択部  
23 映像生成部  
24 仮想空間情報蓄積部  
25 映像生成タイミング発生部  
26 映像選択生成部  
27 多地点音声入力部  
28 音声合成部  
31 多地点の会議端末からの多方向映像  
32 各会議端末へ送信する会議空間の映像  
33 多地点の会議端末からの制御情報  
34 多地点の会議端末からの音声  
35 各会議端末へ送信する音像  
361 各会議端末1で受信する多方向映像



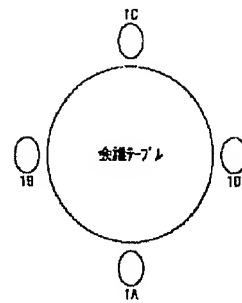
【図1】



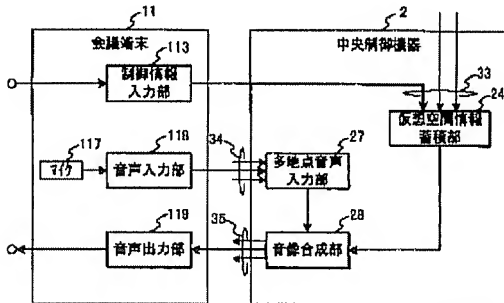
【図2】



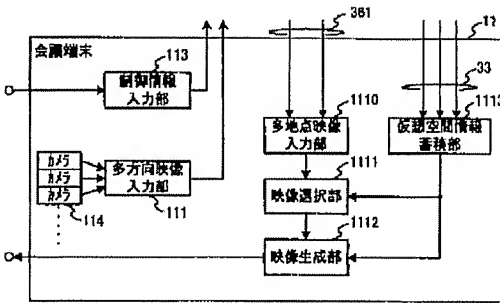
【図6】



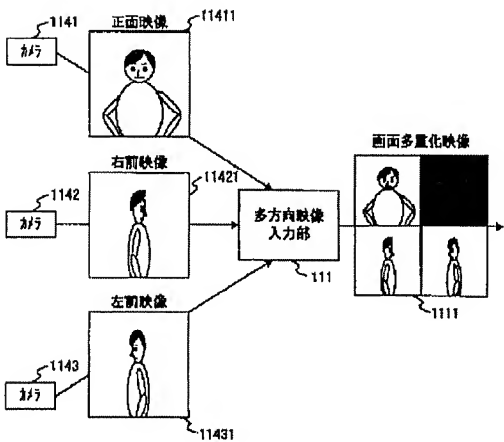
【図3】



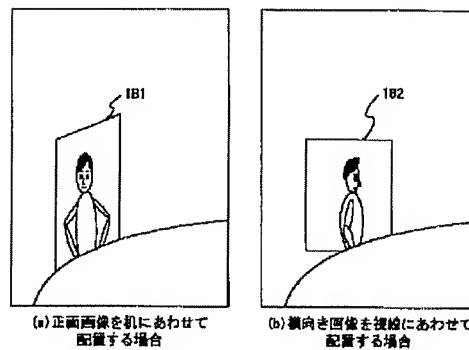
【図4】



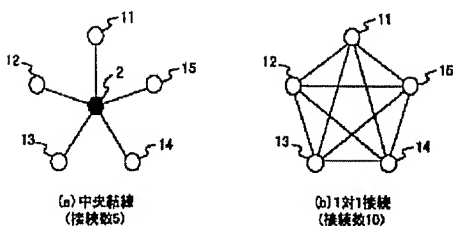
【図5】



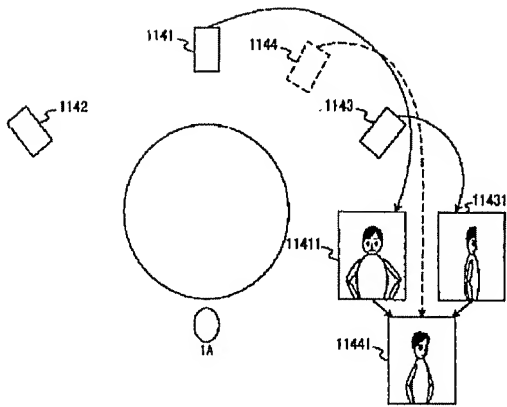
【図7】



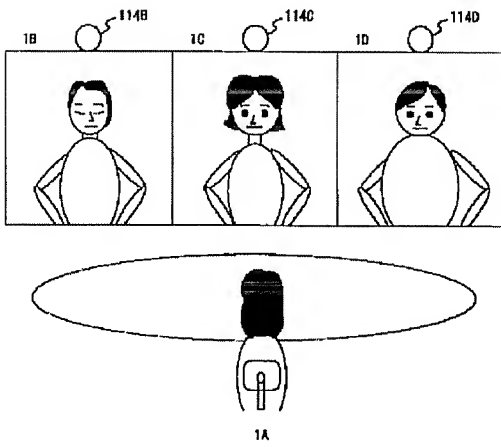
【図9】



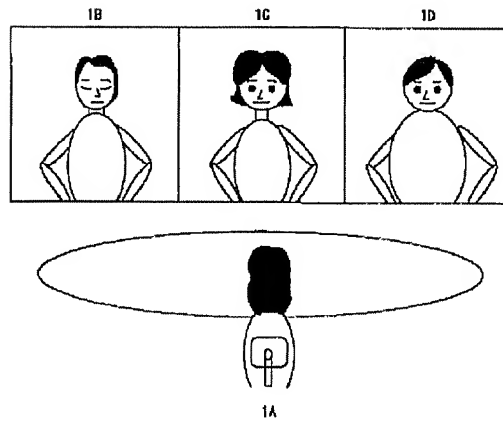
【図8】



【図11】



【図10】



【図12】

